

Nachrichtenwerte und computerunterstützte Inhaltsanalyse

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

(1978). Nachrichtenwerte und computerunterstützte Inhaltsanalyse. *ZUMA Nachrichten*, 2(2), 3-11. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-210749>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

NACHRICHTENWERTE UND COMPUTERUNTERSTÜTZTE INHALTSANALYSE

Neben der Aufgabe, sozialwissenschaftliche Untersuchungen konzeptionell und technisch zu unterstützen, betreut ZUMA noch eine Reihe sogenannter "Grundlagenforschungsprojekte" - Versuchsreihen, die das Ziel verfolgen, Methoden und Instrumente der empirischen Sozialforschung weiterzuentwickeln, neue Verfahren zu erproben und schließlich in möglichst standardisierter Form dem Benutzer zugänglich zu machen.

Eines der Felder, auf denen ZUMA solche Entwicklungsarbeit leistet, ist die Inhaltsanalyse. Sie stellt sich immer stärker als eine der wichtigsten Methoden der Sozialwissenschaft heraus: Die "klassische" Dokumentenanalyse wird mehr und mehr ergänzt durch Studien über sozialen Wandel, Interaktionsanalysen und nicht zuletzt durch den Einsatz systematischer Kategorisierungsverfahren in der Verschlüsselung von Antworten auf offene Fragen im sozialwissenschaftlichen Interview. Ralf Lisch und Jürgen Kriz leiten daraus den Schluß ab, daß die Inhaltsanalyse "die wesentliche - wenn auch oft implizite - Grundlage anderer Modelle sozialwissenschaftlicher Realitätserfassung" ist.¹⁾

Die Arbeit, die ZUMA im Bereich inhaltsanalytischer Grundlagenforschung leistet, umfaßt drei Bereiche:

1. die Konstruktion und Prüfung standardisierter Kategorienschemata,
2. Verfahren der Reliabilitätssteigerung in der konventionellen Inhaltsanalyse und schließlich
3. die Entwicklung und Erprobung von computerunterstützten Verfahren der Kategorisierung von Kommunikationsinhalten.

Einen der Forschungsschwerpunkte in diesem dritten Bereich stellt der folgende Beitrag vor.

Die computerunterstützte Inhaltsanalyse - oft mitleidig als nette, aber teure Spielerei ohne großen praktischen Nutzen belächelt - zeigt bei ihrem Einsatz

¹⁾ Ralf Lisch, Jürgen Kriz: Grundlagen und Modelle der Inhaltsanalyse.

Reinbek b. Hamburg: Rowohlt 1978, S. 31.

im ZUMA Vorzüge, die sich nicht zuletzt in Mark und Pfennig niederschlagen. Die ZUMA-Vercodungsabteilung verschlüsselt z.B. die "offenen" Antworten zur Berufstätigkeit in Umfragen¹⁾ inzwischen standardmäßig elektronisch. Der konventionellen Vercodung dieser Angaben lag ein umfangreiches Kategoriensystem zugrunde, das sehr hohe Anforderungen an das Erinnerungsvermögen der menschlichen Vercoder stellte (291 Kategorien mußten präsent sein). Andererseits aber war das Schema aufgrund der leichtverständlichen Regeln der Zuordnung von Berufsangabe und Code intellektuell derart trivial, daß Routinisierungseffekte nicht ausblieben. Die Berufsverschlüsselung bot sich aus diesen Gründen als ein ideales Anwendungsfeld für den Computer an. Die Ergebnisse können sich sehen lassen: Bei etwa gleichen Kosten wie für die konventionelle Vercodung erhält der Forscher zusätzlich zu einer Verschlüsselung mit etwa 99% "richtigen" Zuordnungen einen maschinenlesbaren Textdatensatz der verschlüsselten Antworten, der ihm einen leichten Zugriff auf die Originaldaten ermöglicht. Darüberhinaus ist der Verschlüsselungsvorgang exakt rekonstruierbar (welcher Text hat welchen Code?) und dadurch auf Wunsch bequem zu modifizieren - bis hin zur Verwendung eines anderen Codesystems.

Alle diese Vorzüge haben dazu geführt, computerunterstützte Inhaltsanalysen auch einmal auf andere Texte anzuwenden als auf die Angaben zur Berufstätigkeit. Für sie liegt ja zugegebenermaßen der Einwand auf der Hand, es handele sich dabei nicht gerade um ein typisches Anwendungsfeld der Inhaltsanalyse im landläufigen Sinne. Aus diesem Grund haben wir als Testbeispiel einer Untersuchung, die den herkömmlichen Vorstellungen einer "echten" Inhaltsanalyse besser entspricht, eine sog. "Nachrichtenwertstudie" gewählt. In ihr geht es darum, Merkmale herauszufinden, die der von den Massenmedien dargestellten Realität eigen sind, oder anders ausgedrückt: Welche Merkmale eines "Ereignisses" bewirken, daß Journalisten 1. überhaupt darüber berichten und ihm 2. besondere Aufmerksamkeit widmen?

¹⁾ Vgl. ZUMANACHRICHTEN 1, S. 18 f.

Eine solche Studie bot sich für unseren Zweck an, weil sie einen Bereich der Sozialwissenschaften berührt, in dem die Inhaltsanalyse das sicher wichtigste Auswertungsverfahren ist, nämlich die Massenkommunikationsforschung. Außerdem sind gerade Nachrichtenwerte seit einiger Zeit zu einem zentralen Untersuchungsgegenstand der Publizistikwissenschaft geworden. Es geht dabei um Faktoren, die Journalisten zu erhöhter Aufmerksamkeit veranlassen: Kulturelle Nähe eines Ereignisses, die Prominenz beteiligter Personen, Betroffenheit von seinen Auswirkungen, ein einfach strukturierter Ablauf und Konfliktrichtigkeit sind einige davon.

Um zu prüfen, welchen Beitrag die computerunterstützte Inhaltsanalyse auf diesem Forschungsfeld leisten kann, haben wir die lokale Berichterstattung dreier Tageszeitungen ("Mannheimer Morgen", "Rhein-Neckar-Zeitung", "Rheinpfalz") über Mannheim im Sommer 1977 vierzehn Tage lang gesammelt und maschinenlesbar abgeschrieben - Grundvoraussetzung für eine Verarbeitung mit dem Computer. Dabei entstand ein Textdatensatz aus 753 Artikeln mit 9.194 Sätzen und etwa 1.100.000 Zeichen. Diese Beiträge hat ein Forschungsseminar an der Universität Mannheim zunächst konventionell verschlüsselt, d.h. eine Reihe von Ereignismerkmalen nach einem zuvor ausgearbeiteten Kategoriensystem dem Text zugeordnet. Diese Vercodung kann damit als eine Art Vergleichsmaßstab für die Ergebnisse des elektronischen Verfahrens dienen.

Bei der Auswahl der Kategorien für die computerunterstützte Inhaltsanalyse war allerdings zu berücksichtigen, daß nicht alle Kategorien eines Verschlüsselungsschemas gleichermaßen zur Bearbeitung mit dem Computer geeignet sind. Das hängt mit einer Eigentümlichkeit der elektronischen Inhaltsanalyse zusammen: Menschliche Vercoder interpretieren die in der Regel verbalen Umschreibungen einer Kategorie aufgrund ihres Sprachverständnisses und Vorwissens so, daß sie konkrete Textelemente selbständig als zu ihr gehörig definieren. (Und das Problem der Verschlüsselungsreliabilität entsteht in dem Maße, in dem dieses Vermögen uneinheitlich ausgebildet ist und die Zuordnungen unterschiedlich ausfallen). Der Computer andererseits ist zu

solchen Leistungen gar nicht erst in der Lage (und hat deshalb auch keine Reliabilitätsprobleme). Ihm muß der Forscher etwa Begriffe, die eine Kategorie hinreichend definieren, explizit, in Form einer operationalen Definition vorgeben. Das elektronische Inhaltsanalyse-System TEXTPACK, das bei ZUMA seit einiger Zeit implementiert ist und ständig weiterentwickelt wird, ist darüber hinaus der Beschränkung unterworfen, daß - zumindest noch gegenwärtig - nur sog. Einwort-Analysen möglich sind, bei denen jeweils nur einzelne Wörter den Kategorien zugeordnet werden. Das hat zur Folge, daß die maschinelle Bearbeitung nur bei Codes möglich ist, die sich mit Hilfe einzelner Wörter hinreichend definieren lassen. Das ist z.B. nicht der Fall, wenn man Texte nach der Dauer der in ihnen geschilderten Vorgänge klassifizieren will - ein nicht unwesentliches Merkmal, das den Nachrichtenwert eines Ereignisses bestimmt.

Einige der Faktoren jedoch, die den Journalisten ein Ereignis stärker beachten lassen, können wir verhältnismäßig einfach und zeitsparend auch mit Hilfe der eingeschränkten Möglichkeiten des verwendeten Programmsystems elektronisch analysieren. Ein Beispiel dafür ist die Berichterstattung über kriminelle Handlungen. Es läßt sich hier etwa die Hypothese prüfen, daß Zeitungen über kriminelle Handlungen nicht proportional zu deren Häufigkeit berichten (und damit z.B. nicht über Ladendiebstähle häufiger als über Morde), sondern daß sie ihre Informationshäufigkeit nach dem Grad der Lebensbedrohung eines Verbrechens gewichten. Da die Umgangssprache mit Hilfe einzelner Wörter recht genau zwischen kriminellen Handlungen unterscheiden kann, die in unterschiedlichem Ausmaß lebensbedrohliche Auswirkungen haben (etwa "Diebstahl", "Raub" und "Totschlag", "Mord"), ist es möglich, dem Computer eine Liste von Wörtern anzugeben - ein "Wörterbuch krimineller Handlungen" sozusagen -, die er, wenn er sie im Text findet, bestimmten Kategorien zuordnen kann. Ein vorläufiges Wörterbuch für dieses Ereignismerkmal umfaßt bereits 257 Bezeichnungen für kriminelle Handlungen. Es wurde am Text selbst entwickelt, d.h. ein alphabetisches Wörterverzeichnis aus den zugrundeliegenden Zeitungsbeiträgen (mit Hilfe von TEXTPACK

leicht herzustellen) wurde auf alle Begriffe durchsucht, die in mehr oder weniger eindeutiger Weise eine strafbare Handlung bezeichnen. Die dabei ausgewählten Wörter haben wir vier Kategorien zugeordnet:

100 Kriminelle Handlung allgemein (ohne Angabe des Schweregrads)

101 Verbrechen mit Lebensbedrohung (Mord, Totschlag, schwere Körperverletzung usw.)

102 Verbrechen ohne Lebensbedrohung (Sittlichkeitsdelikte, Raub, Einbruch usw.)

103 Vergehen und Ordnungswidrigkeiten (Betrug, Urkundenfälschung, Diebstahl, Verkehrsdelikte usw.)

Das Programm erlaubt zusätzlich eine Kennzeichnung von solchen Wörtern, die den für sie gewählten Code nicht ganz eindeutig erfüllen und deshalb vorsichtshalber noch einmal im Kontext angeschaut werden sollten. Die nachstehende Tabelle zeigt einen Auszug aus dem "Wörterbuch krimineller Handlungen" (S. 10 f).

Die dreistellige Zahl vor jedem Wort gibt an, in welche Kategorie es bei seinem Vorkommen eingeordnet werden soll. Worte, die zusätzlich mit einer Eins markiert sind, sind bezogen auf ihre Kategorie ambivalent und bedürfen einer Überprüfung im Textzusammenhang.

Das Wörterbuch ist noch vorläufig, und es wird sofort deutlich, daß mit Hilfe der Wörter noch viel feinere Kategorisierungen möglich wären als die, die von uns mit Rücksicht auf die parallele konventionelle Vercodung festgesetzt wurden. Zusätzliche Unterscheidungen zwischen Betrugs- und Verkehrsdelikten, zwischen Verbrechen mit und ohne Sachschaden usw. wären ohne weiteres angebbar. Das gegenwärtige vierteilige Kategoriensystem ist aber bereits differenziert genug, um die angeführte Hypothese über den Nachrichtenwert krimineller Handlungen zu bestätigen: Es zeigte sich, daß bezogen auf die untersuchten Zeitungsartikel tatsächlich über solche kriminellen Akte häufiger berichtet wurde, die in hohem Ausmaß lebensbedrohliche Auswirkungen haben, obwohl angenommen werden muß, daß sie objektiv

seltener sind als solche ohne lebensgefährdenden Charakter.

Dieser Gesamtbefund läßt sich differenzieren, indem die drei Zeitungen getrennt untersucht werden. Es ergeben sich bedeutsame Unterschiede in den relativen Häufigkeiten, mit denen die vier Kategorien besetzt waren, und in der Ausführlichkeit der Nachrichten. Dabei zeigt sich ein nicht zu unterschätzender Vorteil der elektronischen Inhaltsanalyse: Es ist völlig unproblematisch, die Zähleinheiten bei der Analyse zu variieren. Durch eine einfache Parameterangabe kann der Computer entweder Sätze, Beiträge, einzelne Zeitungsexemplare oder beliebige andere vom Forscher definierte Einheiten auf das Vorkommen von Kategorien überprüfen.

Ein weiteres in Arbeit befindliches Wörterbuch ist ein Orts- und Straßenlexikon der Stadt Mannheim. Das Wörterbuch kann Orts- und Straßennamen insgesamt 78 statistischen Bezirken und Unterbezirken zuordnen. Dadurch haben wir auf relativ einfache Weise die Möglichkeit, Aussagen über Art und Umfang der Repräsentation von Stadtbezirken mit bestimmter Sozialstruktur in den Lokalteilen der drei Zeitungen zu machen. Auch dies ist eine Dimension, die den Nachrichtenwert eines Ereignisses bestimmt. Technisch ist die Möglichkeit für derartige stadtgeographische und sozialstrukturelle Analysen mit dem bei ZUMA entwickelten Filterprogramm TEXTPACK-OSIRIS-INTERFACE gegeben: Das Programm ist z.B. in der Lage, systematisch alle Texte auszuwählen, in denen zuvor definierte Codes (etwa alle Ortsbezeichnungen eines Stadtteils) vorkommen.

Die Perspektiven, die sich aus diesen noch relativ unkomplizierten Anwendungen elektronischer Inhaltsanalyse ergeben, lassen sich leicht ausmalen. Sie werden in dem Maße konkret, in dem es gelingt, relevante Wörterbücher zu konstruieren. Dabei ist nicht so sehr die Definition von Kategorien problematisch, sondern vielmehr die Festlegung von Zuordnungsregeln, die

dem Computer in Form der vorstrukturierten Wörterbücher eingegeben werden. Dieses Problem ist natürlich auch bei einer konventionellen Verschlüsselung durch menschliche Vercoder gegenwärtig: Die Subsumtion eines Wortes, das verschlüsselt werden soll, unter eine bestimmte Kategorie erfolgt ja vor dem Hintergrund des Verständnisses von Alltagssprache, und dieses Verständnis ist immer mehr oder weniger eindeutig, ausgeprägt und allgemeingültig. Man kann deswegen nicht davon ausgehen, daß es in allen Fällen zu unumwundenen Ja-Nein-Entscheidungen kommt: Menschliche Vercoder verschlüsseln mit unterschiedlicher subjektiver Wahrscheinlichkeit. Bei der maschinellen Inhaltsanalyse besteht die Gefahr, dieser Ambivalenz unseres lebendigen Sprachverständnisses Gewalt anzutun und durch Apriori-Definitionen von Zuordnungsregeln auszuschalten.

Um die elektronische Inhaltsanalyse in diesem Sinne lebensnaher und zugleich flexibler zu machen, wird ZUMA langfristig dazu übergehen, die Wörterbücher für bestimmte Themenbereiche so zu konstruieren, daß die Zuordnungen zu den Kategorien auf Stufen unterschiedlicher Sicherheit vorgenommen werden können. Voraussetzung dafür sind empirische Untersuchungen über das Verhalten von professionellen Vercodern. Ihre subjektive Wahrscheinlichkeit, mit der sie Textmaterial bestimmten Kategorien unterordnen, muß skaliert werden, um so jedem Wort eines Wörterbuchs einen Index beizugeben, der ausdrückt, mit welcher Sicherheit menschliche Vercoder die Zuordnung dieses Wortes zu einer der möglichen Kategorien vornehmen. Methodisch ist für die Skalierungen dieser subjektiven Wahrscheinlichkeiten Verhältnisniveau anzustreben, wenn die Indizes über die Menge der Wörter und Kategorien vergleichbar sein sollen. Bei der praktischen Arbeit mit dem Inhaltsanalysensystem ist dann die Möglichkeit gegeben, Kategorisierungen auf unterschiedlichen Eindeutigkeitsniveaus auszuführen bzw. Verschlüsselungen, die sich umgangssprachlich nicht nachvollziehen lassen, von der theoretischen Arbeit auszuschließen.

Anfragen über die Anwendung elektronischer Inhaltsanalyse bei ZUMA richten Sie bitte an Hans D. Klingemann, Klaus Schönbach oder Bernd Wegener.

Auszug aus dem "Wörterbuch krimineller Handlungen"

(A - M)

100	1	Alkoholeinwirkung	101	Blutvergießen
101		Amokfahrt	102	Brandstiftung
100		Angeklagte	103	1 Demoliert
100		Angeklagten	103	Dieb
100		Angeklagter	103	Diebe
100	1	Angetrunkenener	103	Diebestour
100	1	Angriff	103	Diebin
100	1	Angriffsmöglichkeit	103	Diebstähle
100		Anklage	103	Diebstahl
100	1	Anzeige	103	Diebstahlabsicht
100	1	Anzeigen	103	Diebstahls
102	1	Aufgebrochen	100	1 Durchsucht
102	1	Aufzubrechen	100	1 Durchwühlt
102		Autoknacker	102	1 Eingedrungen
102		Automatenaufbrecher	102	1 Eingeschlagen
102		Automatenaufbrüche	100	1 Einsätze
102		Automatenknacker	100	1 Einsatz
102		Baustelleneinbruch	100	1 Einsatzes
101	1	Beckenfraktur	100	1 Einsatzfahrt
100	1	Beckenverletzung	100	1 Einsatzort
100	1	Beschädigt	103	1 Entriß
100	1	Beschädigte	103	1 Entrissen
103		Bestohlen	103	Entwendet
103		Bestohlene	103	Entwendeten
100		Beute	103	1 Entwich
100	1	Beweis	100	Erbeutet
101	1	Bewußtlos	100	Erbeutete
101	1	Bewußtlosigkeit	100	Erbeuteten
100		Blutalkoholtest	101	Ermordet
100		Blutprobe	101	Ermordete

ZUMA

103	1	Ertappt	103		Hausdetektiven
100		Fahndung	101	1	Heimtückischen
103	1	Fahrerlaubnis	100	1	Hilferufe
100		Festgenommen	100	1	Hilflos
100		Festgenommenen	103	1	Höchstgeschwindigkeit
100		Freiheitsstrafe	100	1	Justizvollzugsanstalt
100	1	Funkstreife	102	1	Kellerbrand
100	1	Funkstreifenwagens	101	1	Knochenreste
100		Gefängnis	100	1	Körperprellungen
100	1	Geschädigte	100	1	Körperverletzungen
100	1	Geschädigten	100	1	Kopfplatzwunde
101	1	Geschossen	100	1	Kopfplatzwunden
103		Gestohlen	100	1	Kopfprellungen
101	1	Getötet	100	1	Kopfverletzungen
100		Gewalt	103		Kreditbetrag
100		Gewalttätigkeiten	100	1	Kriminalbeamte
100		Gewaltanwendung	100		Kriminalität
100		Gewaltsam	100	1	Kriminalpolizei
100		Gewaltsamen	100	1	Kripo
100	1	Gezerrt	100	1	Landgericht
103		Grabschändung	100	1	Landgerichts
100		Grausame	101	1	Lebensgefahr
100		Grausamkeit	101	1	Leiche
100	1	Großeinsatz	102	1	Leichtverletzte
100		Hafrichter	101		Messerheld
100	1	Haschisch	100		Mindeststrafe
100	1	Hauptkommissar	101		Mord